34

MECHANISMS AND ENGINEERING SCIENCE

Dingmar van Eck

1. Introduction

Use of "mechanism talk" is ubiquitous in engineering science (e.g., Chandrasekaran and Josephson 2000; Goel 2013).[1] Philosophical discussions of mechanisms also frequently invoke engineered systems, such as pumps, car engines, mouse traps, toilets, soda vending machines, and the like in illustrating various aspects of mechanisms and mechanistic explanation (see Levy 2014). Nevertheless, focused philosophical analyses of the structure of mechanistic explanations in engineering science are scarce (see van Eck 2015). Reference to engineered systems in discussions of mechanisms and mechanistic explanation is often a loose metaphor, not a conceptualization that offers sophisticated understanding of what mechanistic explanation looks like in engineering practice. Moreover, philosophical work that aims to elucidate the connection(s) between engineering and systems biology—connections that practicing engineers and biologists have been stressing for more than a decade (e.g. Csete and Doyle 2002)—is also few and far between, in particular with respect to the use of engineering principles in the construction of mechanistic explanations in systems biology (see Braillard 2015). In this chapter I address both these issues.

In this chapter I give an outline of the structure of mechanistic explanation in engineering science, and organize this discussion around two features that extend the mechanistic program toward explanation when applied to engineering science. First, in section 2, I show that in engineering, two distinct sub-types of role function—"behavior function" and "effect function"—are employed in the functional individuation of mechanisms, rather than role function *simpliciter*. Empirically informed understanding of mechanistic explanation in engineering science requires sensitivity to this distinction (van Eck 2015). I illustrate this point in terms of reverse engineering and malfunction explanations in engineering science.

Second, in section 3, I discuss connections between (control) engineering and systems biology, focusing on the usage of engineering principles in the construction of mechanistic explanations in systems biology. Systems biology has adopted engineering tools and principles, in particular from control engineering, to model and explain complex biological systems. These tools are often in the service of characterizing the organization of mechanisms in abstract, truncated fashion. I discuss a case of heat shock response in *Escherichia coli* to illustrate the role of engineering principles in mechanistic explanation in systems biology (see El-Samad et al. 2005; Braillard 2015).

This case again shows the relevance of distinguishing behavior from effect function and, moreover, gives means to elaborate a key issue in a recent and general discussion on the explanatory power of mechanistic explanations, viz. to flesh out the distinctions between the explanatory desiderata of "completeness and specificity" (Craver 2007) and "abstraction" (Levy and Bechtel 2013). Rather than being in competition, as some authors have it, I argue that these desiderata are suitable for different explanation-seeking contexts.

2. Mechanistic explanation in engineering science

By now, quite a few accounts of mechanistic explanation are on offer in the literature. Although they come in different flavors, there is broad consensus on a number of key features:

> All mechanistic explanations begin with (a) the identification of a phenomenon or some phenomena to be explained, (b) proceed by decomposition into the entities and activities relevant to the phenomenon, and (c) give the organization of entities and activities by which they produce the phenomenon.
>
> *(Illari and Williamson 2012: 123; see also Chapter 1)*

Mechanistic explanations thus explain how mechanisms, i.e. organized collections of entities and activities, produce phenomena (Machamer et al. 2000; Glennan 2005; Bechtel and Abrahamsen 2005; Craver 2007).

Role function ascription is considered crucial for (b) decomposition and (c) the elucidation of mechanisms' organization (Machamer et al. 2000; Craver 2001; Illari and Williamson 2010; see Chapter 8). As Machamer et al. (2000) write:

> Mechanisms are identified and individuated by the activities, and entities that constitute them, by their start and finish conditions, and by their functional roles. Functions are the roles played by entities and activities in a mechanism. To see an activity as a function is to see it as a component in some mechanism, that is, to see it in a context that is taken to be important, vital, or otherwise significant
>
> *(Machamer et al. 2000: 6)*

Mechanistic role functions thus refer to activities that make a contribution to the workings of mechanisms of which they are a part, and mechanistic organization is key for the ascription of functions. For instance, in the context of explaining the circulatory system's activity of "delivering goods to tissues," the heart's "pumping blood through the circulatory system" is ascribed a function relative to organizational features such as the availability of blood, and the manner in which veins and arteries are spatially organized (Craver 2001: 64).

This perspective on the general structure of mechanistic explanation, and the importance of (role) functional individuation of mechanisms, finds widespread support in the literature on mechanistic explanation in the life sciences. Frequently, in this literature, mechanisms of technical artifacts, such as clocks, mousetraps, and car engines, are invoked as metaphors to elucidate features of biological mechanisms (Craver 2001) and features of mechanisms in general (Glennan 2005; Darden 2006; Illari and Williamson 2012). The mechanistic concept of role function, and its utility in the functional individuation of mechanisms, has likewise been explicated in terms of mechanisms of technical artifacts such as car engines (Craver 2001).

However, as mentioned in section 1, reference to such technical mechanisms must not be understood as providing insight into mechanistic explanation in engineering science per se since engineers use multiple notions of function in the functional individuation of technical mechanisms, rather than the concept of role function *simpliciter* (van Eck 2015). (And, as we will see, the distinction between "contextual" and "isolated" descriptions of an entity's activity (Craver 2001) also does not capture the distinction in engineering concepts of function.)

Function is a key term in engineering and an ambiguous one (e.g. Chandrasekaran and Josephson 2000). A variety of function notions are used in mechanism individuation and explanation in engineering science, and the precise notion of function invoked depends on the explanatory and design task at hand. Contrary to explanation in other sciences, explanation in engineering science cannot be seen in isolation from design. For instance, failure analysis—malfunction explanation—is an important type of explanation (Bell et al. 2007) that has as its ultimate aim the improvement of technical systems; also reverse engineering explanation is not "merely" mechanistic explanation since its ultimate aim is the redesign and subsequent improvement of extant technical systems (Otto and Wood 2001). Similarly, explanation is in the service of conceptual design in knowledge base-assisted designing in which (mechanistic) explanations of the workings of extant technical systems and their components are archived and put to use to develop novel design specifications (Stone and Wood 2000). Below I zoom in on two such contexts and the relevance of different notions of engineering function for the functional individuation of technical mechanisms in these contexts, viz. reverse engineering explanation and malfunction explanation. As we will see, specific notions of function are optimally "engineered" for specific explanatory settings (I focus here on explanatory contexts, not the design contexts to which they are related).

Function has no uniform meaning in engineering: different approaches advance different conceptualizations (Erden et al. 2008), and some researchers use the term with more than one meaning simultaneously (Chandrasekaran and Josephson 2000). This ambiguity led to philosophical analysis of the precise meanings of function involved. Vermaas (2009) regimented the spectrum of available function meanings into three "archetypical" engineering conceptualizations of function: *behavior function*—function as the desired behavior of a technical artifact; *effect function*—function as the desired effect of behavior of a technical artifact; and *purpose function*—function as the purpose for which a technical artifact is designed.[2] In the ensuing discussion on reverse engineering explanation and malfunction explanation, the notions of behavior function and effect function are most relevant.

Behavior functions are typically modeled as conversions of flows of materials, energy, and signals, where input flows and output flows in the conversion (are assumed to) match in terms of physical conservation laws (Stone and Wood 2000; Otto and Wood 2001). For instance, the function "loosen/tighten screws" of an electric screwdriver is then represented as a conversion of input flows of "screws" and "electricity" into corresponding output flows of "screws," "torque," "heat," and "noise" (see Stone and Wood 2000: 364). Since these descriptions of functions are specified such that input and output flows match in terms of physical conservation laws—here, the conservation of energy through the conversion of electrical energy into rotational, thermal, and acoustic energy—they are taken to refer to specific physical behaviors of technical artifacts (Vermaas 2009).

Effect function descriptions refer to only the technologically relevant *effects* of the physical behaviors of technical artifacts: the requirements are dropped that descriptions of these effects meet conservation laws and that matching input and output flows are specified (Vermaas 2009). The function of an electric screwdriver is then described simply as, say, "loosen/tighten screws," leaving the physical antecedents of this effect unmentioned. Behavior function descriptions thus refer to the "complete" behaviors involved, including features like thermal and acoustic energy flows, whereas effect functions refer to subsets of these behaviors, i.e. desired effects.[3]

Engineering descriptions and explanations of the workings of extant technical artifacts and artifact designs are often constructed by functionally decomposing functions into a number of sub-functions. The relationships between functions and sets of their sub-functions are often graphically represented in functional decomposition models. Like the concept of function, such models come in a variety of "archetypical" flavors (van Eck 2011). In the context of reverse engineering explanation and malfunction explanation, the relevant ones are *behavior functional decomposition*—a model of an organized set of behavior functions, and *effect functional decomposition*—a model of an organized set of effect functions.

In reverse engineering explanations, elaborate behavior functions and functional decompositions are used; in malfunction explanations, less detailed effect functions and functional decompositions are employed.

In engineering science, reverse engineering and engineering design go hand in glove (e.g. Otto and Wood 2001; Stone and Wood 2000). Consider Otto and Wood's (2001) reverse engineering and redesign method, in which a reverse engineering phase in which reverse engineering explanations are developed for existing artifacts precedes and drives a subsequent redesign phase of those artifacts. The goal of the reverse engineering phase is to explain how existing artifacts produce their overall (behavior) functions in terms of underlying mechanisms, i.e. organized components and sub-functions (behaviors) by which overall (behavior) functions are produced. These explanations are subsequently used in the redesign phase to identify components that function sub-optimally and to either improve them or replace them with better-functioning ones.

In the reverse engineering phase, an artifact is first broken down component-by-component, and hypotheses are formulated concerning the functions of those components. In this method, functions are behavior functions and are represented by conversions of flows of materials, energy, and signals. Since the aim of the reverse engineering phase is to understand in detail the manner in which an extant technical system operates, elaborate behavior function descriptions are used. Descriptions of input–output conversions give more relevant details than descriptions of effects. After this analysis, a different reverse engineering analysis commences in which components are removed, one at a time, and the effects are assessed of removing single components on the overall functioning of the artifact. Such single-component removals are used to detail the behavior functions of the (removed) components further. The idea behind this latter analysis is to compare the results from the first and second reverse engineering analysis to gain a potentially more nuanced understanding of the functions of the components of the (reverse engineered) artifact. Using these two reverse engineering analyses, a behavior functional decomposition of the artifact is then constructed in which the behavior functions of the components are specified and interconnected by their input and output flows of materials, energy, and signals

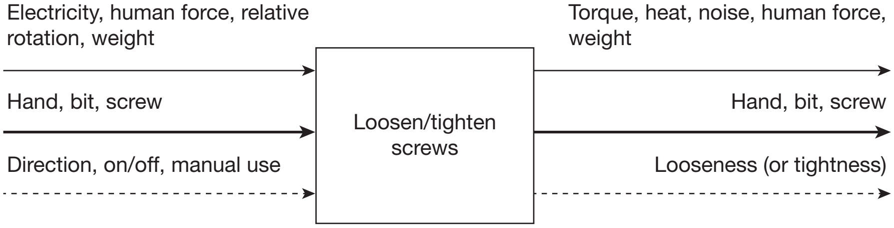

Figure 34.1 Overall behavior function of an electric power screwdriver. Thin arrows represent energy flows; thick arrows represent material flows, dashed arrows represent signal flows (adapted from Stone and Wood 2000: 363, figure 2)

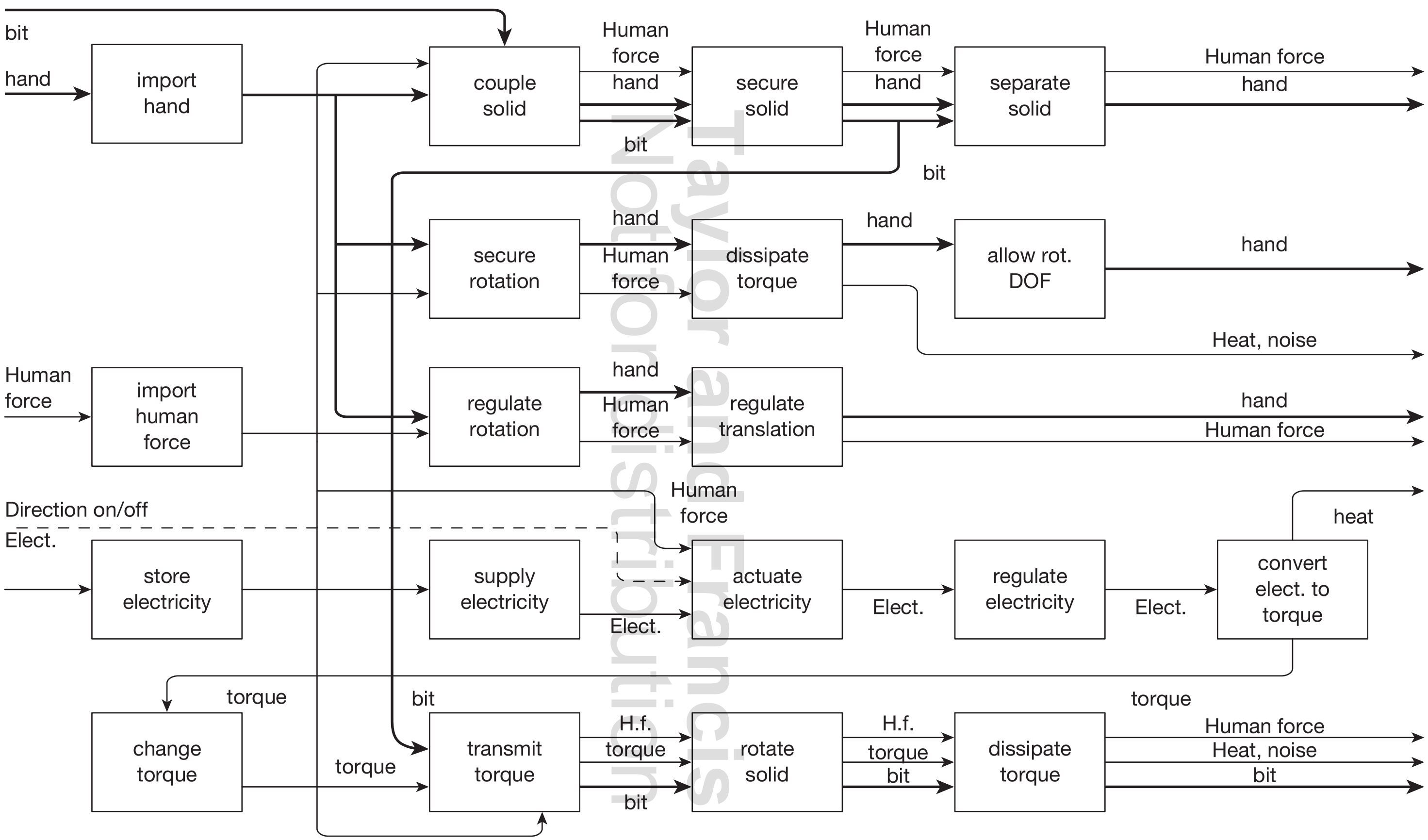

Figure 34.2 Behavior functional decomposition of an electric power screwdriver. Thin arrows represent energy flows; thick arrows represent material flows, dashed arrows represent signal flows (adapted from Stone and Wood 2000: 364, figure 4)

(Otto and Wood 2001). Such models represent parts of the mechanisms by which technical systems operate, to wit: causally connected behaviors of components. Examples of an overall behavior function and behavior functional decomposition of a reverse engineered electric screwdriver are given in Figures 34.1 and 34.2, respectively.

In the model in Figure 34.2, temporally organized and interconnected behaviors are described. Components of artifacts are described in Otto and Wood's method in tables, what in engineering are called "bills of materials," together with a model, called an "exploded view," of the components composing the artifacts. Taken together, these component and behavior functional decomposition models provide functional individuations and representations of mechanisms of artifacts.

Such (behavior functional decomposition) models are subsequently used to identify sub-optimally functioning components and so drive succeeding redesign phases (Otto and Wood 2001). The focus here is on the reverse engineering explanation part of the methodology.

In malfunction explanation, this detail in mechanistic models is, however, not required: engineers take it that less detailed effect functions and functional decompositions there do a better explanatory job.

In malfunction analysis, explanation-seeking questions of the following format arise:

Why does artifact x not serve the expected function to ϕ?

Such questions are *contrastive*: why malfunction, rather than normal function? In the engineering literature, malfunction explanations that answer contrastive questions list different and fewer mechanistic features than reverse engineering explanations which answer questions about normal behavior or function.[4] Such explanations are constructed using effect functions and functional decompositions.

Malfunction explanations in engineering pick out only a few features of mechanisms, i.e. those causal factors—failing components or sub-mechanisms—that are taken to make a difference to the occurrence of a specific malfunction, as well as some coarse-grained details of the containing mechanism to understand where the fault is located. Yet most information about structural and behavioral specifics of malfunctioning components/sub-mechanisms, and their containing mechanisms, is left out (Hawkins and Woollons 1998; Bell et al. 2007).[5,6]

Consider, by way of example, the Functional Interpretation Language (FIL) methodology for malfunction analysis and explanation (Bell et al. 2007). In FIL, functions are effect functions and represented in terms of their *triggers* and *effects*. Triggers describe input states that actuate physical behaviors which result in certain (expected) effects. So triggers are the input conditions for effects, i.e. functions, to be achieved but they are not the immediate physical antecedents of these effects. These physical antecedents, i.e. input flows, are not referred to in trigger-effect descriptions and neither do these descriptions meet conservation laws. For instance, consider the function description "depress_brake_pedal"—"red_stop_lamps_lit" of a car's stop light (Bell et al. 2007: 400), in which electromagnetic radiation ("light") is "created" rather than being converted from other energies. This description, rather, is a summary of some salient features of (manipulating) such artifacts; depressing the brake pedal will, if the system functions properly, result in the lighting of the stop lamps, i.e. the effect function.

According to Bell et al. (2007), such trigger and effect representations serve two explanatory ends in malfunction analyses: first, they *highlight* relevant behavioral features of a given artifact, i.e. effects, and, simultaneously, provide the means to *ignore* less relevant or irrelevant behavioral features, i.e. physical behaviors underlying these effects; second, they support assessing which components are malfunctioning (Bell et al. 2007: 400–1).

For instance, the trigger-effect representation "depress_brake_pedal"—"red_stop_lamps_lit" highlights the input condition of a pedal being depressed, and the resulting desired effect of lighted lamps, yet ignores the structural and behavioral specifics of the brake pedal and stop lamps, such as the pedal lever and electrical circuit mechanisms, as well as the energy conversions—e.g., mechanical energy conversions into electricity—that are needed to achieve this effect. Such representations only highlight those features that are considered explanatorily relevant to assess malfunctioning systems, and omit reference to physical behaviors/energy conversions by which the desired effects are achieved.

Second, such trigger-effect descriptions support comparing normally functioning technical systems with malfunctioning ones (Bell et al. 2007). Trigger-effect descriptions support assessing whether the expected effects in fact obtain, and, if not, which and how components are malfunctioning (Bell et al. 2007). A normally functioning artifact, say the car's stop lights, has both a trigger and an effect occurring; the brake pedal is depressed and the stop lights are lit. Trigger-effect descriptions support analysis of two varieties of malfunction. First, a trigger may occur, yet fail to result in the intended effect. Say, the brake pedal is depressed, yet the stoplights are not on. Second, a trigger may not be occurring, yet the effect is nevertheless present. Say, the brake pedal is not depressed, yet the stoplights are on (see Bell et al. 2007). Such analysis of the actual states of triggers and effects allows one to focus on the most likely causes of failure (Bell et al. 2007). Say, if the pedal is depressed and the lights fail to ignite, the first likely causes to investigate may be whether the electrical circuits in the lights are broken or the "on/off" connection between the brake and electrical circuitry (connected to the lamp) is damaged. On the other hand, if the pedal is not depressed and the lights are lit, a first likely cause to investigate may be whether the "on/off "connection between the brake and the electrical circuitry is damaged. To support more detailed malfunction analyses, functions are often decomposed into sub-functions in FIL. An example of a functional decomposition of a two-ring cooking hotplate is given in Figure 34.3.

The usage of effect functions and functional decompositions in FIL is the optimal choice given that function descriptions are used to black-box or suppress reference to unwanted behavioral and structural details. Effect function descriptions only highlight the relevant difference-making properties with respect to malfunctioning artifacts, whereas more elaborate behavior function descriptions include irrelevant details such as, say, the thermal energy generated when lamps are lit.

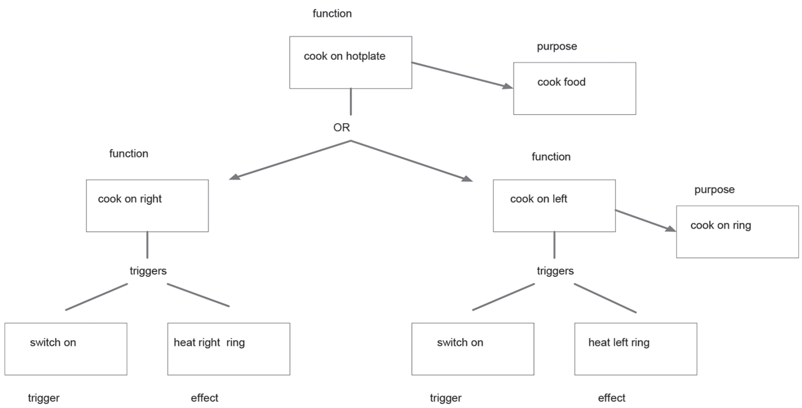

Figure 34.3 Effect functional decomposition of a two-ring cooking hotplate (adapted from Bell et al. 2007)

The upshot of these two cases is that explanations in engineering are furnished relative to explanatory objectives and, importantly, the level of detail included in these explanations hinges on specific concepts of technical function. Engineering scientists simplify or increase the details of explanations—functional decompositions—depending on the explanatory purpose at hand, and these adjustments are made using specific concepts of technical function (compare e.g. Figures 34.2 and 34.3). In reverse engineering explanation, elaborate or "complete" descriptions of mechanisms are provided, in terms of behavior functions and functional decompositions, to answer the question of how a technical system exhibits a given overall behavior. In malfunction explanation, less elaborate "sketches" of mechanisms are provided in terms of effect functions and functional decompositions, referring only to some mechanistic features, namely those difference-making factors that mark the *contrast* between normal functioning and malfunctioning technical systems. So, depending upon explanatory context, mechanisms are individuated in different ways using different conceptualizations of function in engineering science. Neither function conceptualization in itself accommodates both ways in which mechanisms are functionally individuated in engineering science. Behavior and effect function ascriptions are invoked to individuate mechanisms in different ways depending on the task at hand.

However, this distinction in functional individuation, and its reliance on different function concepts, remains opaque when seen from a perspective that conceives of mechanism individuation and mechanistic explanation in terms of mechanistic role function ascription *simpliciter*. The concept of mechanistic role function, an activity that makes a contribution to the workings of a mechanism of which it is a part, admits of two interpretations in the context of engineering science: behavior function on the one hand and effect function on the other. Note that behavior and effect descriptions of function describe, in different ways, the contributions of components to mechanisms of which they are a part. The distinction between behavior and effect function thus is not to be conflated with the distinction between a mechanism description and a description of a mechanism's overall activity. Neither is the behavior-effect function distinction to be conflated with the distinction between "isolated" and "contextual" descriptions of an entity's activity (Craver 2001): isolated descriptions describe activities without taking into account the mechanisms in which they are situated; contextual descriptions describe activities in terms of the mechanistic contexts in which they are situated and to which they contribute. Both behavior and effect functions are of the contextual variety, describing contributions of components to the mechanisms of which they are a part. So to arrive at an empirically informed understanding of explanatory practices in engineering, and at consistency of the general structure of mechanistic explanation with these practices, regimenting the concept of role function into domain-specific engineering concepts of behavior and effect function, i.e. sub-types of role function, is needed.

I now turn to another facet of the relationship between mechanistic explanation and engineering that has received little sustained analysis: the usage of engineering principles in the construction of mechanistic explanations in systems biology. Here we see again the relevance of knowing the ways in which mechanisms are functionally individuated in engineering: the manner in which biological mechanisms are individuated in engineering terms also hinges on specific *engineering* conceptualizations of function. To better understand the specifics of mechanism individuation along engineering lines in systems biology, sensitivity to the varieties of engineering function and functional individuation is thus called for. In the case I discuss below, mechanism individuation hinges on the use of *effect* function descriptions and ascriptions.

3. Explanation and systems thinking: where engineering and systems biology meet

Although philosophy, it seems, is only recently picking up on the fruitful cross-talk between engineering and systems biology (see Braillard 2015), with Wimsatt's (e.g. 2006) work being a notable exception, engineers and systems biologists alike have been stressing the conceptual ties for more than a decade (Hartwell et al. 1999; Lazebnik 2002). With biological data about complex biological systems exploding during the last 20 years or so because of (functional) genomics projects and the like, opportunities to understand complex biological systems in far greater detail became available. Yet cashing out that promise also signaled the need for new tools that enabled massive data analysis and integration to build explanatory models of these complex systems with a scale and complexity hitherto unknown (see Chapter 27). Here is where, amongst others, engineering tools came in.

As Hartwell et al. (1999), for instance, commented with respect to engineering representational schemes:

> In our opinion, the most effective language to describe functional modules and their interactions [in systems biology] will be derived from the synthetic sciences, such as computer science or engineering, in which function appears naturally.
>
> *(Hartwell et al. 1999: C49)*

Or as Csete and Doyle (2002) commented with respect to systemic organization:

> Advanced technologies [like cars and airplanes] and biology have extremely different physical implementations, but they are far more alike in systems-level organization than is widely appreciated
>
> *(Csete and Doyle 2002: 1664)*

Functional engineering parlance and systemic organization come together in the connection between control engineering and systems biology, in which (effect functional) decomposition and control principles governing (the construction of) engineering systems are used to characterize complex biological systems:

> Some insightful recent papers advocate a similar modular decomposition of biological systems according to the well defined functional parts used in engineering and, specifically, engineering control theory.
>
> *(Tomlin and Axelrod 2005: 4219)*

A case in point is research by El-Samad et al. on the mechanism(s) to counter heat shock in *Escherichia coli* (El-Samad et al. 2005; see Tomlin and Axelrod 2005; Braillard 2015). Heat shock response is a widely conserved response of cells to cope with environmental stress brought about by unusual increases in temperature, involving the induced expression of heat shock proteins. Such temperature increases can damage proteins by breaking down their tertiary structures. Heat shock proteins come in two varieties and mitigate this effect in two different ways: molecular chaperones do so by refolding denatured proteins and proteases by degrading denatured proteins. If the response is sufficiently swift and massive, cell death can be prevented by protein repair and/or removal of damaged proteins. The response needs to be tightly controlled in the sense that it is only activated in the case of heat shock, since the response is highly energy-consuming and would make too high energy demands if heat shock proteins would be produced all the time. Cells thus must maintain a delicate balance between the protective effect of heat shock protein production and the metabolic cost of overproducing these proteins. In *E. coli*, the RNA polymerase cofactor ø^{32} promotes the transcription of heat shock proteins. After heat shock stress—temperature increase—ø^{32} activity increases, resulting in the transcription of specific heat shock gene promoters which initiate the transcription of genes that in turn encode specific heat shock proteins—chaperones and proteases. This heat shock protein expression, when appropriate, prevents cell death. This mechanism uses both feed-forward and feedback loops that process information about temperature and the folding state of proteins in the cell. ø^{32} activity is crucial in all this and depends on a feed-forward mechanism that senses temperature and controls ø^{32} transcription, and feedback regulatory mechanisms that register the folding levels of proteins (levels of denatured cellular protein) and degrade ø^{32}. These regulatory feedback mechanisms are crucial to ensure that ø^{32} synthesis, activity, and stability are brought back to normal levels after a sufficient number of heat shock proteins have been produced and the threat to cell death is averted.

The above qualitative information on the heat shock response system is well known. El-Samad and his group (2005) went further and constructed a quantitative, mathematical model of the heat shock response to "use this description to pose questions about the regulatory architecture of the system" (El-Samad et al. 2005: 2737), i.e. the dynamical, mechanistic organization that sustains the heat shock response. They came up with an elaborate mathematical model consisting of 31 equations and seven parameters. Now, to make the model computationally tractable and pose and answer questions about the dynamical, mechanistic organization of the system, the original model had to be trimmed down. This model reduction was effected by various simplifications and, importantly, the salient modularity of the system which made it possible to decompose the system into functional modules, described in terms of effect functions, along control engineering lines. The resulting model was taken to be a "simplified yet reasonably accurate version of the original model" (El-Samad et al. 2005: 2737).

As Braillard (2015) stressed, control engineering principles played an important heuristic role in this model reduction, and thus in the discovery of the mechanism's core organizational features that sub-serve its overall regulatory behavior. The close analogy between engineered systems and biological ones with respect to functional modular organizations sub-serving regulatory processes makes this possible. As El-Samad et al. (2005) explain:

> Control and dynamical systems theory is a discipline that uses modular decompositions extensively to make modeling and model reduction more tractable. Because biological networks are themselves complex regulation systems, it is reasonable to expect that seeking similarities with the functional modules traditionally identified in engineering schemes can be particularly useful.
>
> *(El-Samad et al. 2005: 2737)*

In control engineering, decomposition into functional modules (modules defined in terms of their effect-role functions) often begins with identification of the process to be regulated called the "plant" (see Lind 1994), for instance altitude regulation of an airplane or temperature regulation of a thermostat. Modules of the system that contribute to the regulation are described in terms of their contributing functions, the most common of which are "sensors," "detectors," "controllers," "actuators," and "feed-forward" and "feedback" signals. For instance, in a simple heating system, the plant is the temperature regulation process, which is achieved, inter alia, by a sensor module which measures ambient temperature, calculates the deviation from the desired temperature, and feeds this information into the thermostat (controller). The thermostat then outputs signals that are sent to an actuator (heat fuel valve) that generates an actuation signal (e.g. fuel to furnace) that corrects deviation from the desired temperature. The sensor module again measures the ambient temperature and, if needed, feeds back information on temperature deviations to the controller, and so on.

El-Samad et al. (2005) applied this control engineering perspective to the *E. coli* heat shock response system. In this application, the protein folding task (the refolding of denatured proteins) is taken to be the process to be regulated (plant), the feed-forward signal (sent by a sensor) is the temperature-dependent translational efficiency of ø^{32} synthesis, the controller is the level of ø^{32} activity, chaperones function as actuators of the plant (the actuated plant input is the number of molecular chaperones), and sensors measure plant output (the amount of denatured protein), which in turn is fed back to the controller.

This decomposition allowed El-Samad et al. (2005) to construct a simplified model consisting of just six equations and 11 parameters in which each equation describes the behavior of a module. They remark:

> This model provides useful insight into the heat shock system design architecture. It also suggests a mathematical and conceptual modular decomposition that defines the functional blocks or submodules of the heat shock system. This decomposition is drawn by analogy to manmade control systems and is found to constitute a canonical blueprint representation for the heat shock network.
>
> *(El-Samad et al. 2005: 2736)*

What we here thus see is that analogical reasoning with respect to regulation processes, and the functional architecture sub-serving these processes in engineered and biological systems, led to a functional modular decomposition of a biological system that laid bare core organizational features of the system by which it produces regulatory behavior. Engineering tools here serve as a discovery heuristic for a mechanism's core organizational features that sub-serve its overall regulatory behavior (see Braillard 2015; Chapter 19, this volume). This usefulness of engineering concepts, i.e. modular decompositions in terms of effect functions, is not specific to the *E. coli* case, and generalizes to a variety of cases (see Tomlin and Axelrod 2005) and suggests a general discovery heuristic:

> If the heat shock mechanism can be described and understood in terms of engineering control principles, it will surely be informative to apply these principles to a broad array of cellular regulatory mechanisms and thereby reveal the control architecture under which they operate.
>
> *(Tomlin and Axelrod 2005: 4220)*[7]

In concluding this chapter, I suggest that this case gives relevant insights into a general discussion on explanatory power in the recent mechanisms literature by providing an empirical illustration of the complementarity of two allegedly competing perspectives on the explanatory power of mechanistic explanations.

I have argued elsewhere that differences between two main (allegedly) competing perspectives on the explanatory power of mechanistic explanations, "completeness and specificity" (Craver 2007) and "abstraction" (Levy and Bechtel 2013), essentially boil down to differences in the notions of difference making endorsed in these accounts and that they are in fact not in competition (van Eck 2015). They are rather suitable for different explanation-seeking contexts. Whereas abstraction dictates that mechanistic explanations should only list the "primary factors" responsible for the occurrence of system function, "completeness and specificity" prescribes that in addition to primary ones, "higher-order factors" should also be described, which concern factors influencing the precise manner in which a system function occurs or those sub-serving the primary factors. The *E. coli* case gives an empirical illustration of this view.

The notion of "robustness" looms large in the *E. coli* case, as well as in systems biology and engineering in general. Robust systems—ones resilient to perturbations to parts of the mechanism or the environment in which it functions—require complex sub-systems dedicated to counteracting perturbations (Kitano 2004). This holds both for complex biological systems and (most) engineered systems. Think, for instance, of all the sub-systems of an airplane dedicated to counteracting changes to make it fly in the appropriate manner, or the sub-systems in *E. coli* that play a role in counteracting the effects of heat shock on protein deformation—chaperons and proteases. As El-Samad et al. (2005) elaborate:

> The modular decomposition of the hsr [heat shock response] shows a level of complexity not justified by the basic functionality demanded from an operational heat shock system. A simple and operational heat-shock system would consist solely of a temperature sensor . . . and a transcriptional/translational apparatus that responds appropriately to temperature changes.
>
> *(El-Samad 2015: 2738)*

Why, then, is additional complexity present? Computational modeling indicated that "complexity is indeed necessary to achieve robustness, noise rejection, speed of response, and economical use of cellular resources, much like engineering systems" (El-Samad 2015: 2738).

Complexity and robustness provide an interesting slant on "abstraction" and "completeness and specificity." Depending on the questions one asks with respect to complex, robust systems, either "completeness and specificity" or "abstraction" are better suited. For instance, one may address the question: "How does the mechanism (execute its) regulatory function?", or the joint questions: "How does the mechanism (execute its) regulatory function?" and "Why does it (execute its) function in a robust manner?" If one is interested in the key organizational details that enable complex systems to function, abstract description suffices. If, on the other hand, one is also interested in the mechanistic details that enable a mechanism to function in a robust fashion, more specific and elaborate descriptions are called for. In the latter case, one is not only interested in the "primary factors" responsible for the occurrence of system function, but also in the "higher-order factors" influencing the precise manner in which it occurs or those sub-serving the primary factors (see Weisberg 2007).

To round up, the functional individuation of mechanisms—in terms of behavior and effect function ascription and decomposition—proceeds differently in engineering science than the manner in which it is taken to work in the life sciences. Understanding these specifics is required to understand the structure of mechanistic explanation in engineering science and, moreover, adds to our understanding of the ways in which tools and insights from engineering are used in mechanism individuation and explanation in systems biology. Finally, cases from engineering and systems biology give general insights into the explanatory power of mechanistic models in specific explanation-seeking contexts.

Notes

1 It is a pleasure thanking Phyllis Illari and Stuart Glennan for the opportunity to write this chapter and for their helpful comments on a previous version.

2 The term "archetypical" here refers to "most common"; the three conceptualizations of function are not meant to be exhaustive. For instance, some engineers use "function" to refer to intentional behaviors of agents (see van Eck 2010). In reverse engineering analyses, "function" refers to actual or expected behavior, without the normative connotation "desired."

3 Behavior and effect functions thus have a partly common semantic structure: certain aspects or features of behaviors that they both refer to. They are dissimilar in the sense that behavior function descriptions refer to additional behavioral aspects, not referred to in effect function descriptions, so as to make these descriptions accord with physical conservation laws. The relation between behavior and effect function is asymmetrical in the sense that effects, being subsets of behaviors, are straightforwardly derivable from behaviors, but not vice versa. From a given effect one cannot automatically derive the behavior of which the effect is a part. Cars that run on gas operate by means of different energy conversions than cars that run on electricity, yet both display the same effects; say, delivering acceleration. The semantic structure that they partly have in common creates the possibility and need to be pluralist about mechanistic role functions, i.e. different ways to conceive of the role functions of mechanisms, in the context of engineering science. I defend this pluralism about mechanistic role functions later on in this section. To be sure, I am thus not advocating a pluralist view about functions of mechanisms with a completely dissimilar semantic structure.

4 Reverse engineering explanation, like mechanistic explanation in general, is not contrastive, whereas malfunction explanation is. The role of contrasts, essential to counterfactual accounts of explanation, seems not vital to most accounts of mechanistic explanation. Mechanistic explanations are often taken to track mechanisms that exhibit productive continuity, and are typically not construed in counterfactual fashion. Counterfactual reasoning, rather, is often invoked in analyses of mechanism discovery and in explanatory relevance assessments where interventions on putative components are stressed (Craver 2007). Malfunction explanations thus make for an interesting extension of mechanistic conceptions of explanation, since they are both mechanistic and contrastive.

5 That is, structural and behavioral characteristics are considered irrelevant in a first-round functional analysis of malfunction. After this analysis, more detailed behavioral models of components and their behaviors are used for identifying specific explanatorily relevant structural and behavioral characteristics of malfunctioning components/sub-mechanisms (Bell et al. 2007). However, immediately specifying these details in functional models is taken to result in listing a lot of irrelevant details.

6 Malfunction explanations in engineering thus exemplify Garson's (2013) "functional sense of mechanism" (see Chapter 8); a malfunction is seen as a breakdown of a mechanism, not as the result of a specific mechanism for malfunction.

7 The analysis I gave in this section illustrates what Glennan and Illari call "methodological mechanism" (see Chapter 1): seen from an epistemological and methodological perspective, biology and engineering have much in common and tools and insights from the latter can help address explanatory issues in the former (and vice versa). I thus do not address the metaphysical nature of mechanisms and how differences play out in this regard between biology and engineering. So, for instance, whether biological mechanisms are modular like (most) technical ones or whether we impose modularity on them to understand how they work is, although a very intriguing question, not one I am concerned with here. Neither do I focus here on possible differences that may emerge between the life sciences and engineering when we consider normative conceptions of function. It might transpire, however, that these differences are not so great as some might suspect, at least in the context of explanation: of course, engineers design and build technical systems-mechanisms with desired ("proper") functions, yet a "functional sense of mechanism" (Garson 2013; see Chapter 8) suffices to account for a token mechanism with a function that it fails to perform, both in engineering and biological contexts. Furthermore, in engineering contexts of explanation, engineering sub-types of role function do the explanatory work, not an etiological conception of function (van Eck and Weber 2014).

References

Bechtel, W., and A. Abrahamsen. (2005). “Explanation: A Mechanist Alternative”, *Studies in History and Philosophy of Biological and Biomedical Sciences* 36:421–41.

Bell, J., Snooke, N., and C. Price. (2007). “A Language for Functional Interpretation of Model Based Simulation”, *Advanced Engineering Informatics* 21:398–409.

Braillard, P.A. (2015). “Prospects and Limits of Explaining Biological Systems in Engineering Terms”. In: *Explanation in Biology* (Eds. P.A. Braillard & C. Malaterre), Springer, 319–44.

Chandrasekaran, B., and J.R. Josephson. (2000). “Function in Device Representation”, *Engineering with Computers* 16:162–77.

Craver, C.F. (2001). “Role Functions, Mechanisms, and Hierarchy”, *Philosophy of Science* 68:53–74.

Craver, C.F. (2007). *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*. New York: Oxford University Press.

Csete, M.E., and J.C. Doyle. (2002). “Reverse Engineering of Biological Complexity”, *Science* 295: 1664–9.

Darden, L. (2006). *Reasoning in Biological Discoveries*. Cambridge: Cambridge University Press.

El-Samad, H., Kurata, H., Doyle, J.C., Gross, C.A., and M. Khammash. (2005). “Surviving Heat Shock: Control Strategies for Robustness and Performance”, *PNAS* 102(8):736–41.

Erden, M.S., Komoto, H., Van Beek, T.J., D’Amelio, V., Echavarria, E., and T. Tomiyama. (2008). “A Review of Function Modeling: Approaches and Applications”, *Artificial Intelligence for Engineering Design, Analysis, and Manufacturing* 22:147–69.

Garson, J. (2013). “The Functional Sense of Mechanism”, *Philosophy of Science* 80:317–33.

Glennan, S. (2005). “Modeling mechanisms”, *Studies in the History and Philosophy of the Biological and Biomedical Sciences* 36(2):375–88.

Goel, A.K. (2013). “A 30-Year Case Study and 15 Principles: Implications of an Artificial Intelligence Methodology for Functional Modeling”, *AIEDAM* 27(3):203–15.

Hartwell, L.H., Hopfield, J.J., Leibner, S., and A.W. Murray. (1999). “From Molecular to Modular Cell Biology”, *Nature* 402:C47–C52.

Hawkins, P.G., and D.J. Woollons. (1998). “Failure Modes and Effects Analysis of Complex Engineering Systems using Functional Models”, *Artificial Intelligence in Engineering* 12(4):375–97.

Illari, P., and J. Williamson. (2010). “Function and Organization: Comparing the Mechanisms of Protein Synthesis and Natural Selection”, *Studies in History and Philosophy of Biological and Biomedical Sciences* 41:279–91.

Illari, P., and J. Williamson. (2012). “What is a Mechanism? Thinking About Mechanisms Across the Sciences”, *European Journal for Philosophy of Science* 2:119–35.

Kitano, H. (2004). “Biological Robustness”, *Nature* 5:826–37.

Lazebnik, Y. (2002). “Can a Biologist Fix a Radio?—Or, What I Learned While Studying Apoptosis”, *Cancer Cell* 2:179–82.

Levy, A. (2014). “Machine-Likeness and Explanation by Decomposition”, *Philosopher’s Imprint* 6:1–15.

Levy, A. and W. Bechtel. (2013). “Abstraction and the Organization of Mechanisms,” *Philosophy of Science* 80:241–61.

Lind, M. (1994). “Modeling Goals and Functions of Complex Industrial Plants”, *Applied Artificial Intelligence* 8:259–83.

Machamer, P.K., Darden, L., and C.F. Craver. (2000). “Thinking About Mechanisms”, *Philosophy of Science* 57:1–25.

Otto, K.N., and K.L. Wood. (2001). *Product Design: Techniques in Reverse Engineering and New Product Development*. Upper Saddle River, NJ: Prentice Hall.

Stone, R.B., and K.L. Wood. (2000). “Development of a Functional Basis for Design”, *Journal of Mechanical Design* 122:359–70.

Tomlin, C.J., and Axelrod, J.D. (2005). “Understanding Biology by Reverse Engineering the Control”, *PNAS* 102(12):4219–20.

Van Eck, D. (2010). “On the Conversion of Functional Models: Bridging Differences Between Functional Taxonomies in the Modeling of User Actions”, *Research in Engineering Design* 21(2):99–111.

Van Eck, D. (2011). “Supporting Design Knowledge Exchange by Converting Models of Functional Decomposition”, *Journal of Engineering Design* 22(11–12):839–58.

Van Eck, D. (2015). "Mechanistic Explanation in Engineering Science", *European Journal for Philosophy of Science* 5(3):349–75.

Van Eck, D., and E. Weber. (2014). "Function Ascription and Explanation: Elaborating an Explanatory Utility Desideratum for Ascriptions of Technical Functions", *Erkenntnis* 79:1367–89.

Vermaas, P.E. (2009). "The Flexible Meaning of Function in Engineering", *Proceedings of the 17th International Conference on Engineering Design (ICED 09)* 2:113–24.

Weisberg, M. (2007). "Three Kinds of Idealization", *The Journal of Philosophy* 104(12):639–59.

Wimsatt, W. (2006). *Re-Engineering Philosophy for Limited Beings: Piecewise Approximations to Reality*. Cambridge, MA: Harvard University Press.